Integrating Current Analyses of the Breast Cancer Microbiome

Sidra Sohail
ssohail@luc.edu

Follow this and additional works at: https://ecommons.luc.edu/grs

https://ecommons.luc.edu/grs/2021/Posters/4

This Restricted is brought to you for free and open access by the Conferences at Loyola eCommons. It has been accepted for inclusion in Graduate Research Symposium by an authorized administrator of Loyola eCommons. For more information, please contact ecommons@luc.edu.

This work is licensed under a Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 License.
Integrating Current Analyses of the Breast Cancer Microbiome

Sidra Sohail, Michael Burns, PhD
Loyola University Chicago

Abstract

This project integrates all available studies related to breast cancer and the mammary microbiome to 1) reassess the original findings in light of advances in this rapidly progressing field and 2) incorporate all the data available as a large meta-analysis to identify general trends and specific differences across patient cohorts and studies. In order to reassess original findings of each study, the 16S rRNA sequence data will be retrieved from SRA and analyzed using a bioinformatics pipeline, which will use an amplicon sequence variant. The results from the bioinformatics pipeline will be compared with the findings from the original studies. In order to identify specific trends and differences across patient cohorts and studies, I will compile the results from each of the available studies into a single dataset to assess for trends in microbial community interactions with breast cancer across studies. As different patient cohorts, 16S rRNA variable regions, and experimental protocols were used across the studies of interest, I will need to statistically account for these covariates in my model. The final ASV tables from the bioinformatics pipeline analysis for each study will be merged and assessed using multivariate analysis.

Background & Research Questions

There are two questions our research will address:

1. How robust are the original findings to re-analysis with modern tools?
2. Are there new biological insights about the disease that emerge from a combined meta-analysis of all available studies?

Research Approach

The microbiome is the collection of bacteria that reside in tissues, and in cancer it has been implicated in a variety of tissues [2]. The nature of it is such that there are specific microbial effects in carcinogenicity for the different tissues [2]. In terms of the breast tissue, it was previously thought that it was sterile; however, it is now known that breast tissue has a diverse microbial community where differences have been observed between healthy and tumor tissue [1; 2; 3]. Currently, in microbiome research, taxonomic databases have been massively expanded and cleaned allowing for more accurate and updated taxonomic assignment, where pipelines and databases account for unforeseen biases. This project will analyze data from four studies investigating the microbiota of breast tissue, which include the papers by Urbaniak et al. [1], Hieken et al. [2], Xuan et al. [3], and Chan et al. [4]. Currently, I am in the process of analyzing data from the Urbaniak paper and Hieken paper.

There are two questions our research will address:

1. How robust are the original findings to re-analysis with modern tools?
2. Are there new biological insights about the disease that emerge from a combined meta-analysis of all available studies?

Recent Findings, Next Steps, & Acknowledgements

Our current results show that our findings are different from the original studies. The bacterial abundance from our findings from the Hieken[2] and Urbaniak[1] data analysis differ from what is reported in the original respective papers. It is exciting to discover that our preliminary findings do not align with the original results, as this means that there are biological factors in the updated datasets (SILVA) that no one has explored yet. The next steps are to analyze the Xuan[3] and Chan[4] data and run them through DADA2 and phyloseq, and implement an approach similar to what is stated in the Research Approach section. I would like to thank Dr. Burns for his great mentorship, and guidance in implementing code. Generous funding was provided by the Burns Lab.

References


Figure a): An abundance bar plotting bacterial composition at the genus level from the DADA2 analysis of the Hieken data, which includes Aphagea, Bacillus, Choanozoa, Pectobacterium, Serratia, Streptococcus, and Streptophyta.

Figure b): An abundance bar plotting bacterial composition at the genus level from the DADA2 analysis of the Urbaniaik data, which includes Escherichia-Shigella, Flavobacterium, Lachnospiraceae NK3A21 group, Mesorhizobium, and Salmonella.

Figure c): An abundance bar plot from the original Hieken study showing bacterial composition of the benign and cancer samples at the Phylum, Family, and Genus levels.

Figure d): An abundance bar plot from the DADA2 analysis of the Hieken data showing bacterial abundance of the benign and cancer samples at the Phylum, Family, and Genus levels. The bacterial taxa in Figures c) & d) do not overlap at any taxonomic level, and do not contain any similarities.